HAUSDORFF MOMENT PROBLEM VIA FRACTIONAL MOMENTS

1. Introduction

In Applied Sciences a variety of problems, formulated in terms of linear boundary values or integral equations, leads to a Hausdorff moment problem. Such a problem arises when a given sequence of real numbers may be represented as the moments around the origin of non-negative measure, defined on a finite interval, typically $[0,1]$. The underlying density $f(x)$ is unknown, while its moments $\mu_j = \int_0^1 x^j f(x)dx$, $j = 0,1,2,...,$with $\mu_0 = 1$, are known. Next, through a variety of techniques, for practical purposes $f(x)$ is recovered by taking into account only a finite sequence $\{\mu_j\}_{j=0}^M$. Such a process implies that $f(x)$ is well-characterized by its first few moments. On the other hand, it is well known that the moment problem becomes ill-conditioned when the number of moments involved in the reconstruction increases [1,2]. In Hausdorff case, once fixed $(\mu_0, ..., \mu_{M-1})$, the moment μ_M may assume values within the interval $[\mu_M^-, \mu_M^+]$, where [3]

$$\mu_M^+ - \mu_M^- \leq 2^{-2(M-1)} \tag{1.1}$$

If one considers the approximating density $f_M(x) = \exp(-\sum_{j=0}^M \lambda_j x^j)$ by entropy maximization, constrained by the first M moments [4], then its entropy $H[f_M] = -\int_0^1 f_M(x)\ln f_M(x)dx$ satisfies

$$\lim_{\mu_M \to \mu_M^\pm} H[f_M] = -\infty \tag{1.2}$$

Such a relationship is satisfied by any other distribution constrained by the same first M moments, since $f_M(x)$ has maximum entropy. On the other hand $f(x)$ and $f_M(x)$ have the same first M moments and as a consequence, as we illustrate in section 3, the following relationship holds

$$I(f, f_M) =: \int_0^1 f(x) \ln \frac{f(x)}{f_M(x)} dx = H[f_M] - H[f]. \tag{1.3}$$

Here $H[f]$ is the entropy of $f(x)$, while $I(f, f_M)$ is the Kullback-Leibler distance between $f(x)$ and $f_M(x)$.

Equations (1.1)-(1.3) underline once more the ill-conditioned nature of the moment problem. The ill-conditioning may be even enlightened by considering the estimation of the parameters λ_j of $f_M(x)$. The λ_j calculation leads to minimize a proper potential function $\Gamma(\lambda_1, ..., \lambda_M)$[Kesa 4], with

$$\min_{\lambda_1,...,\lambda_M} \Gamma(\lambda_1, ..., \lambda_M) = \min_{\lambda_1,...,\lambda_M} \Big[\ln\Big(\int_0^1 \exp(-\sum_{j=1}^M \lambda_j x^j)dx\Big) + \sum_{j=1}^M \lambda_j \mu_j\Big]. \tag{1.4}$$

$f_M(x)$ satisfies the constraints

$$\mu_j = \int_0^1 x^j \exp(-\sum_{k=0}^M \lambda_k x^k 1)dx, \qquad j = 0, ..., M \tag{1.5}$$

Letting $\mu = (\mu_0, ..., \mu_M)$ and $\lambda = (\lambda_0, ..., \lambda_M)$, (1.5) may be written as the map

$$\mu = \phi(\lambda) \tag{1.6}$$

Then the corresponding Jacobian matrix, which is up to sign a Hankel matrix, has conditioning number $\simeq (1+\sqrt{2})^{4M}/\sqrt{M}$ [5]. All the previous remarks lead to the conclusion that $f(x)$ may be efficiently recovered from moments only if few moments are requested. In other terms, $f(x)$ may be recovered from moments if its information content is spread among first few moments.

In this paper we are looking for a way to overcome the above-quoted difficulties in recovering $f(x)$ from moments. First of all, we assume the infinite sequence of moments $\{\mu_j\}_{j=0}^{\infty}$ to be known. Then, from such a sequence, we calculate fractional moments

$$E(X^{\alpha_j}) =: \int_0^1 x^{\alpha_j} f(x)dx = \sum_{n=0}^{\infty} b_n(\alpha_j)\mu_n, \quad \alpha_j > 0 \tag{1.7}$$

where the explicit analytic espression of $b_n(\alpha_j)$ is given by (2.5). Finally, from a finite number of fractional moments $\{E(X^{\alpha_j})\}_{j=1}^{M}$, we recover $f_M(x) = \exp(-\sum_{j=0}^{M} \lambda_j x^{\alpha_j})$ by entropy maximization [4]. The exponents $\{\alpha_j\}_{j=1}^{M}$ are chosen as follows

$$\{\alpha_j\}_{j=1}^{M}: \quad H[f_M] = \text{minimum} \tag{1.8}$$

The choice of $\{\alpha_j\}_{j=1}^{M}$, according to (1.8), leads to a density $f_M(x)$ having minimum distance from $f(x)$, as stressed by (1.3).
Remark. If the information content of $f(x)$ is shared among first moments, so that ME approximant $f_M(x)$ represents an accurate approximation of $f(x)$, then fractional moments may be accurately calculated by replacing $f(x)$ with $f_M(x)$. As a consequence, function $f_M(x)$ converges in entropy and then in L_1-norm to $f(x)$ [6], and the error obtained replacing $f(x)$ with $f_M(x)$

$$\begin{aligned} \mid E_f(X^{\alpha_j}) - E_{f_M}(X^{\alpha_j}) \mid &\leq \int_0^1 x^{\alpha_j} \mid f(x) - f_M(x) \mid dx \leq \\ &\leq \int_0^1 \mid f(x) - f_M(x) \mid dx \leq \sqrt{2(H[f_M] - H[f])} \end{aligned} \tag{1.9}$$

may be rendered arbitrarily small by increasing M (inequalities in (1.9) are proved in section 3).

2. Fractional moments from moments

Let X a continuous random variable with density $f(x)$ on the support $[0,1]$, with moments of order s, centered in c, $c \in I\!R$

$$\mu_s(c) := I\!E\left[(X-c)^s\right] = \int_0^1 (x-c)^s\, f(x)\, dx, \quad s \in I\!N^* = I\!N \cup \{0\}. \tag{2.1}$$

and moments from the origin $\mu_s =: \mu_s(0)$ related to moments generically centered in c through the relationship

$$\mu_s = \sum_{h=0}^{s} \binom{s}{h} c^{s-h}\, \mu_h(c), \quad s \in I\!N^*. \tag{2.2}$$

It is well known the relationship similar to (2.2) which permits to calculate the (fractional) moment of order $s \in I\!R^+$ (which replaces α_j for notational convenience as in (1.7) and (3.2)) involving all the central moments of a given distribution about the point c.
Firstly, by definition of noncentral moment of order s, we can write $I\!E(X^s) = \int_0^1 x^s f(x)dx$ and then, by Taylor expansion of x^s around c, where $c \in (0,1)$, we have

$$\begin{aligned} x^s &= \sum_{n=0}^{\infty} [x^s]^{(n)}_{x=c}\, \frac{(x-c)^n}{n!} \\ &= \sum_{n=0}^{\infty} \left[\binom{s}{n}\, n!\, x^{s-n}\right]_{x=c} \frac{(x-c)^n}{n!} \\ &= \sum_{n=0}^{\infty} \binom{s}{n} c^{s-n} (x-c)^n \end{aligned} \tag{2.3}$$

where $[k(x)]_{x=c}^{(n)}$ indicates the n-th derivative of the function $k(x)$ wrt x, evaluated at c.
Taking the expectation on both sides of the last equation in (2.3), we get the required relationship

$$\begin{aligned} \mathbb{E}\left(X^s\right) &= \sum_{n=0}^{\infty} \binom{s}{n} c^{s-n} \mathbb{E}\left[(X-c)^n\right] \\ &= \sum_{n=0}^{\infty} b_n \, \mu_n(c) \end{aligned} \tag{2.4}$$

where

$$b_n = \binom{s}{n} c^{s-n}, \quad n \in \mathbb{N}^* \tag{2.5}$$

represents the coefficient of the integral n-order moment of X centered at c.
The formulation of the s-order fractional moments as in (2.4) shows some numerical instabilities which depend on the structure of the relationship between $\mu_n(c)$ and $\mathbb{E}(X^s)$; these instabilities are related to the value of the center c and increase as the order of the central moments becomes high. In particular,

(a) the numerical error $\Delta\mathbb{E}(X-c)^n$ due to the evaluation of $\mathbb{E}(X-c)^n$ in terms of noncentral integral moments $\mathbb{E}(X^h)$, $h \le n$, becomes bigger as c and n increase. In fact,

$$\begin{aligned} |\Delta\mathbb{E}(X-c)^n| &= \left| \sum_{h=0}^{n} (-1)^h \binom{n}{h} c^{n-h} \, \Delta\mathbb{E}(X^h) \right| \\ &\le \sum_{h=0}^{n} \binom{n}{h} c^{n-h} \, |\Delta\mathbb{E}(X^h)| \\ &= \| \, \Delta\, \mathbb{E}(X^h) \, \|_\infty \, \sum_{h=0}^{n} \binom{n}{h} c^{n-h} = \\ &= \| \, \Delta\, \mathbb{E}(X^h) \, \|_\infty \, (1+c)^n \simeq eps\,(1+c)^n, \end{aligned} \tag{2.6}$$

where *eps* corresponds to the error machine.

(b) the numerical error $\Delta\,\mathbb{E}(X^s)$ due to the evaluation of $\mathbb{E}(X^s)$ involving the first M_{max} central moments $\mathbb{E}(X-c)^n$, is given by

$$\begin{aligned} |\Delta\,\mathbb{E}(X^s)| &= \sum_{n=0}^{M_{max}} \binom{s}{n} c^{s-n} \Delta\,\mathbb{E}(X-c)^n \\ &\le \sum_{n=0}^{M_{max}} \left|\binom{s}{n}\right| c^{s-n} \, |\Delta\,\mathbb{E}(X-c)^n| \\ &\le \| \, \Delta\,\mathbb{E}(X-c)^n \, \|_\infty \, c^s \, \max_n \binom{s}{n} \sum_{n=0}^{M_{max}} \left(\frac{1}{c}\right)^n \\ &= \| \, \Delta\,\mathbb{E}(X-c)^n \, \|_\infty c^s \, \max_n \binom{s}{n} \frac{\left(\frac{1}{c}\right)^{M_{max}+1} - 1}{\frac{1}{c} - 1}, \end{aligned} \tag{2.7}$$

with $\max_n \binom{s}{n} = \binom{s}{[s/2]}$ if $[s]$ is even and $\max_n \binom{s}{n} = \binom{s}{[s/2]+1}$ if $[s]$ is odd, where $[x]$ represents the integer part of x. The product of first two factors of the right hand side of (2.7) is an increasing function of c, whilst the last factor gives a function which decreases with c.

Hence, taking in account both (a) and (b), a reasonable choice of c could be $c = \frac{1}{2}$. Further, rewriting the last inequality in (2.7) as

$$|\Delta\, E\!\!\!\!E(X)^s| \leq \|\, \Delta\, E\!\!\!\!E(X-c)^n \,\|_\infty \, c^s \max_n \binom{s}{n} \frac{\left(\frac{1}{c}\right)^{M_{max}+1} - 1}{\frac{1}{c} - 1} < \varepsilon$$

we can reconstruct the s-order fractional moment with a prefixed level of accuracy ε, $\varepsilon > 0$, just involving a number of central moments equal to the value M_{max}.

3. Recovering $f(x)$ from fractional moments

Let be X a positive r.v. on $[0,1]$ with density $f(x)$, Shannon-entropy $H[f] = -\int_0^1 f(x)\ln f(x)dx$ and moments $\{\mu_j\}_{j=0}^\infty$, from which positive fractional moments $E(X^{\alpha_j}) = \sum_{n=0}^\infty b_n(\alpha_j)\mu_n$ may be obtained, as in (2.4)-(2.5).

From [4], we know that the Shannon-entropy maximizing density function $f_M(x)$, which has the same M fractional moments $E(X^{\alpha_j})$, of $f(x)$, $j = 0, ..., M$, is

$$f_M(x) = \exp(-\sum_{j=0}^M \lambda_j x^{\alpha_j}). \tag{3.1}$$

Here $(\lambda_0, ..., \lambda_M)$ are Lagrangean multipliers, which must be supplemented by the condition that the first M fractional moments of $f_M(x)$ coincide with $E(X^{\alpha_j})$, i.e,

$$E(X^{\alpha_j}) = \int_0^1 x^{\alpha_j} f_M(x)dx,\ j = 0, ..., M, \quad \alpha_0 = 1 \tag{3.2}$$

The Shannon entropy $H[f_M]$ of $f_M(x)$ is given as

$$H[f_M] = -\int_0^1 f_M(x)\ln f_M(x)dx = \sum_{j=0}^M \lambda_j E(X^{\alpha_j}). \tag{3.3}$$

Given two probability densities $f(x)$ and $f_M(x)$, there are two well-known measures of the distance between $f(x)$ and $f_M(x)$. Namely the divergence measure $I(f, f_M) = \int_0^1 f(x)\ln\frac{f(x)}{f_M(x)}dx$ and the variation measure $V(f, f_M) = \int_0^1 \mid f_M(x) - f(x) \mid dx$. If $f(x)$ and $f_M(x)$ have the same fractional moments $E(X^{\alpha_j})$, $j = 1, ..., M$ then

$$I(f, f_M) = H[f_M] - H[f] \tag{3.4}$$

holds. In fact $I(f, f_M) = \int_0^1 f(x)\ln\frac{f(x)}{f_M(x)}dx = -H[f] + \sum_{j=0}^M \lambda_j \int_0^1 x^{\alpha_j} f_M(x)dx = -H[f] + \sum_{j=0}^M \lambda_j E(X^{\alpha_j}) = H[f_M] - H[f]$.

In literature, several lower bounds for the divergence measure I based on the variation measure V are available. We shall however use the following bound [7]

$$I \geq \frac{V^2}{2}. \tag{3.5}$$

If $g(x)$ denotes a bounded function, such that $\mid g(x) \mid \leq K$, $K > 0$, by taking into account (3.4) and (3.5), we have

$$\mid E_f(g) - E_{f_M}(g) \mid \leq \int_0^1 \mid g(x) \mid \cdot \mid f(x) - f_M(x) \mid dx \leq K\sqrt{2(H[f_M] - H[f])} \tag{3.6}$$

. Equation (3.6) suggests us what fractional moments have to be chosen

$$\{\alpha_j\}_{j=1}^{M}: \quad H[f_M] = \text{minimum} \tag{3.7}$$

The use of fractional moments in the framework of ME relies on the following two theoretical results. The first is a theorem [8, Th. 2] which guarantees the existence of a probability density from the knowledge of an infinite sequence of fractional moments

Theorem 3.1 [8, Th. 2] If X is a r.v. assuming values from a bounded interval $[0,1]$ and $\{\alpha_j\}_{j=0}^{\infty}$ is an infinite sequence of positive and distinct numbers satisfying $\lim_{j\to\infty}\alpha_j = 0$ and $\sum_{j=0}^{\infty}\alpha_j = +\infty$, then the sequence of moments $\{E(X^{\alpha_j})\}_{j=0}^{\infty}$ characterizes X.

The second concerns the convergence in entropy of $f_M(x)$, where entropy-convergence means $\lim_{M\to\infty} H[f_M] = H[f]$. More precisely,

Theorem 3.2. If $\{\alpha_j\}_{j=0}^{M}$ are equispaced within $[0,1)$, with $\alpha_{M-j+1} = \frac{j}{M+1}$, $j = 0, ..., M$ then the ME approximant converges in entropy to $f(x)$.
Proof. See Appendix.

We just point out that the choice of equispaced points $\alpha_{M-j+1} = \frac{j}{M+1}$, $j = 0, ..., M$ satisfies both conditions of Theorem 3.1, i.e.

$$\lim_{M\to\infty}\alpha_M = 0 \text{ and } \lim_{M\to\infty}\sum_{j=0}^{M}\alpha_j = \lim_{M\to\infty}\frac{1}{M+1}\frac{M}{2}(M+1) = +\infty.$$

As a consequence, if the choice of equispaced α_{M-j+1} guarantees entropy-convergence, then the choice (3.7) guarantees entropy-convergence too.
From a computational point of view, Lagrangean multipliers $(\lambda_1, ..., \lambda_M)$ are obtained by (1.4), and the normalizing constant λ_0 is obtained by imposing that the density integrates to 1. Then the optimal $\{\alpha_j\}_{j=1}^{M}$ exponents are obtained as

$$\{\alpha_j\}_{j=1}^{M}: \quad \min_{\alpha_1,...,\alpha_M}\Big[\min_{\lambda_1,...,\lambda_M}\Gamma(\lambda_1, ..., \lambda_M)\Big]. \tag{3.8}$$

4. Numerical results
We compare fractional and ordinary moments by choosing some probability densities on $[0,1]$.
Example 1. Let be

$$f(x) = \frac{\pi}{2}\sin(\pi x)$$

with $H[f] \simeq -0.144729886$. From $f(x)$ we have ordinary moments satisfying the recursive relationship

$$\mu_n = \frac{1}{2} - \frac{n(n-1)}{\pi^2}\mu_{n-2}, \qquad n = 2, 3, ..., \quad \mu_0 = 1, \quad \mu_1 = \frac{1}{2}.$$

From $\{\mu_n\}_{n=0}^{\infty}$ we calculate $E(X^{\alpha_j}) = \sum_{n=0}^{\infty} b_n(\alpha_j)\mu_n$, as in (2.4)-(2.5). From $\{E(X^{\alpha_j})\}_{j=0}^{M}$ we obtain the ME approximant $f_M(x)$ for increasing values of M, where $\{\alpha_j\}_{j=1}^{M}$ satisfy (3.7).
In Table 1 are reported
a) $H[f_M] - H[f] = I(f, f_M)$ and exponents $\{\alpha_j\}_{j=1}^{M}$ satisfying (3.7), where $H[f_M]$ is obtained using fractional moments.
b) $H[f_M] - H[f] = I(f, f_M)$, where $H[f_M]$ is obtained using ordinary moments.
Inspection of Table 1 allows us to conclude that:

1) Entropy decrease is fast, so that practically 4-5 fractional moments determine $f(x)$.
2) On the converse an high number of ordinary moments are requested for a satisfactory characterization of $f(x)$.
3) Approximately 12 ordinary moments have an effect comparable to 3 fractional moments.
$f(x)$ and $f_M(x)$, obtained by 4-5 fractional moments, are practically indistinguishable.

Table 1
Optimal fractional moments and entropy difference of distributions having an increasing number of common a) fractional moments b) ordinary moments

a)

M	$\{\alpha_j\}_{j=1}^M$	$H[f_M]-H[f]$
1	13.4181	$0.8716E-1$
2	0.00289 4.69275	$0.2938E-2$
3	0.04680 1.84212 13.2143	$0.3038E-3$
4	0.00220 2.76784 13.7293 20.5183	$0.3276E-4$
5	0.0024 2.7000 13.700 20.500 25.200	$0.1016E-4$

b)

M	$H[f_M]-H[f]$
2	$0.9510E-2$
4	$0.2098E-2$
6	$0.7058E-3$
8	$0.4442E-3$
10	$0.3357E-3$
12	$0.3288E-3$

Example 2. This example is borrowed from [9]. Here the authors attempt to recover a non-negative decreasing differentiable function $f(x)$ from the frequency moments ω_n, with

$$\omega_n=\int_0^1 [f(x)]^n dx, \qquad n=1,2,...$$

The authors of [9] realize that other density reconstruction procedures, alternative to ordinary moments, would be desirable. We propose fractional moments density reconstruction procedure. Here

$$f(x)=2\Big[\frac{1}{2}+\frac{1}{10}\ln(\frac{1}{Ax+B}-1)\Big] \qquad B=\frac{1}{1+e^5},\ \ A=\frac{1}{1+e^{-5}}-\frac{1}{1+e^5}$$

with $H[f]\simeq -0.06118227$ ($f(x)$, compared to [9], contains the normalizing constant 2). From $f(x)$ we have ordinary moments μ_n through a numerical procedure. From $\{\mu_n\}_{n=0}^\infty$ we calculate $E(X^{\alpha_j})=\sum_{n=0}^\infty b_n(\alpha_j)\mu_n$, as in (2.4)-(2.5). Finally, from $\{E(X^{\alpha_j})\}_{j=0}^M$ we obtain the ME approximant $f_M(x)$ for increasing values of M, where $\{\alpha_j\}_{j=1}^M$ satisfy (3.7).
Table 2 reports:
a) $H[f_M]-H[f]=I(f,f_M)$ and exponents $\{\alpha_j\}_{j=1}^M$ satisfying (3.7), where $H[f_M]$ is obtained using fractional moments.

b) $H[f_M] - H[f] = I(f, f_M)$, where $H[f_M]$ is obtained using ordinary moments.

Inspection of Table 2 allows us to conclude that:
1) Entropy decrease is fast, so that practically 4 fractional moments determine $f(x)$.
2) An high number of ordinary moments is requested for a satisfactory characterization of $f(x)$.
3) Approximately 14 ordinary moments have an effect comparable to 4 fractional moments.
Functions $f(x)$ and $f_M(x)$, obtained by 4 fractional moments, are practically indistinguishable. As a consequence, we argue that the use of 4 fractional moments is as effective as that of 8 frequency moments (as in [9]). The former ones, indeed, provide an approximant $f_M(x)$ practically indistinguishable from $f(x)$ (see figure 1 of [9]).

Table 2
Optimal fractional moments and entropy difference of distributions having an increasing number of common a) fractional moments b) ordinary moments

a)

M	$\{\alpha_j\}_{j=1}^M$	$H[f_M] - H[f]$
1	1.56280	$0.6278E-2$
2	0.52500	$0.3152E-2$
	3.90000	
3	1.05000	$0.1169E-2$
	3.00000	
	7.87500	
4	0.44062	$0.1025E-3$
	7.65470	
	12.5262	
	63.9093	

b)

M	$H[f_M] - H[f]$
2	$0.5718E-2$
4	$0.1776E-2$
6	$0.1320E-2$
8	$0.6744E-3$
10	$0.3509E-3$
12	$0.2648E-3$
14	$0.1914E-3$

5. Conclusions

In this paper we have faced up the Hausdorff moment problem and we have solved it using a low number of fractional moments, calculated explicitly in terms of given ordinary moments. The approximating density, constrained by few fractional moments, has been obtained by maximum-entropy method. Fractional moments have been chosen by minimizing the entropy of the approximating density. The strategy proposed in the present paper, for recovering a given density function, consists in accelerating the convergence by a proper choice of fractional moments, so obtaining an approximating density by the use of low order moments, as (1.1) suggests.

6. References

[1] D. Fasino, Spectral properties of Hankel matrices and numerical solutions of finite moment problems, J. Comput. Applied Math., **65**, 145-155, (1995).

[2] G. Talenti, Recovering a function from a finite number of moments, Inverse Problems, **3**, 501-517, (1987).

[3] S. Karlin, L.S. Shapley, *Geometry of moment spaces*, AMS Memoirs **12**, Providence RI (1953).

[4] H.K. Kesavan, J.N. Kapur, *Entropy Optimization Principles with Applications*, Academic Press, (1992).

[5] B. Beckermann, The condition number of real Vandermonde, Krylov and positive definite Hankel matrices, Numerische Mathematik, **85**, 553-577, (2000).

[6] J.M. Borwein, A.S. Lewis, Convergence of best entropy estimates, SIAM J. Optimization, **1**, 191-205, (1991).

[7] S. Kullback, A lower bound for discrimination information in terms of variation, IEEE Transaction on Information Theory, IT-13, 126-127, 1967.

[8] G.D. Lin, Characterizations of Distributions via moments, Sankhya: The Indian Journal of Statistics, **54**, Series A, 128-132, 1992.

[9] E. Romera, J.C. Angulo, J.S. Dehesa, The Hausdorff entropic moment problem, J. of Math. Physics, **42**, 2309-2314, (2001).

[10] J.A. Shohat, J.D. Tamarkin, *The problem of moments*, AMS Mathematical Survey, **1**, Providence RI, (1963).

Appendix: Entropy convergence

A.1 Some background

Let's consider a sequence of equispaced points $\alpha_j = \frac{j}{M+1}$, $j = 0, ..., M$ and

$$\mu_j =: E(X^{\alpha_j}) = \int_0^1 t^{\alpha_j} f_M(t)dt, \qquad j = 0, ..., M \tag{A.1}$$

with $f_M(t) = \exp(-\sum_{j=0}^{M} \lambda_j t^{\alpha_j})$. With a simple change of variable $x = t^{\frac{1}{M+1}}$, from (A.1) we have

$$\mu_j = E(X^{\alpha_j}) = \int_0^1 x^j \exp\Big[-(\lambda_0 - \ln(M+1)) - \sum_{j=1}^{M} \lambda_j x^j + M \ln x\Big] dx, \;\; j = 0, ..., M \tag{A.2}$$

which is a reduced Hausdorff moment problem for each fixed M value and a determinate Hausdorff moment problem when $M \to \infty$. Referring to (A.2) the following symmetric definite positive Hankel matrices are considered

$$\Delta_0 = \mu_0, \;\; \Delta_2 = \begin{bmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{bmatrix}, ..., \Delta_{2M} = \begin{bmatrix} \mu_0 & \cdots & \mu_M \\ \vdots & \cdots & \vdots \\ \mu_M & \cdots & \mu_{2M} \end{bmatrix} \tag{A.3}$$

whose (i,j)-th entry $i, j = 0, 1, ...$ holds

$$\mu_{i+j} = \int_0^1 x^{i+j} f_M(x)dx,$$

where $f_M(x) = \exp\Big[-(\lambda_0 - \ln(M+1)) - \sum_{j=1}^{M} \lambda_j x^j + M \ln x\Big]$. The Hausdorff moment problem is determinate and the underlying distribution has a continuous distribution function $F(x)$, with density $f(x)$. Then the massimal mass $\rho(x)$ which can be concentrated at any real point x is equal to zero ([10], Corollary (2.8)). In particular, at $x = 0$ we have

$$0 = \rho(0) = \lim_{i\to\infty} \rho_i^{(0)} =: \frac{|\Delta_{2i}|}{\begin{vmatrix} \mu_2 & \cdots & \mu_{i+1} \\ \vdots & \cdots & \vdots \\ \mu_{i+1} & \cdots & \mu_{2i} \end{vmatrix}} = \lim_{i\to\infty} (\mu_0 - \mu_0^{-(i)}) \tag{A.4}$$

where $\rho_i^{(0)}$ indicates the largest mass which can be concentrated at a given point $x = 0$ by any solution of a reduced moment problem of order $\geq i$ and $\mu_0^{-(i)}$ indicates the minimum value of μ_0 once assigned the first $2i$ moments.

Let's fix $\{\mu_0, ..., \mu_{i-1}, \mu_{i+1}, ..., \mu_M\}$ while only μ_i, $i = 0, ..., M$ varies continuously. From (A.2) we have

$$\Delta_{2M} \cdot \begin{bmatrix} d\lambda_0/d\mu_i \\ \vdots \\ d\lambda_M/d\mu_i \end{bmatrix} = -e_{i+1} \tag{A.5}$$

where e_{i+1} is the canonical unit vector $\in I\!R^{M+1}$, from which

$$0 < \Big[\frac{d\lambda_0}{d\mu_i}, ..., \frac{d\lambda_M}{d\mu_i}\Big] \cdot \Delta_{2M} \cdot \begin{bmatrix} d\lambda_0/d\mu_i \\ \vdots \\ d\lambda_M/d\mu_i \end{bmatrix} = -\Big[\frac{d\lambda_0}{d\mu_i}, ..., \frac{d\lambda_M}{d\mu_i}\Big] e_{i+1} = -\frac{d\lambda_i}{d\mu_i} \qquad \forall i \tag{A.6}$$

A.2 Entropy convergence

The following theorem holds.

Theorem A.1 If $\alpha_j = \frac{j}{M+1}$, $j = 0, ..., M$ and $f_M(x) = \exp(-\sum_{j=0}^{M} \lambda_j x^{\alpha_j})$ then

$$\lim_{M\to\infty} H[f_M] =: -\int_0^1 f_M(x) \ln f_M(x) dx = H[f] =: -\int_0^1 f(x) \ln f(x) dx. \qquad (A.7)$$

Proof. From (A.1) and (A.7) we have

$$H[f_M] = \sum_{j=0}^{M} \lambda_j \mu_j \qquad (A.8)$$

Let's consider (A.8). When only μ_0 varies continuously, taking into account (A.3)-(A.6) and (A.8) we have

$$\frac{d}{d\mu_0} H[f_M] = \sum_{j=0}^{M} \mu_j \frac{d\lambda_j}{d\mu_0} + \lambda_0 = \lambda_0 - 1$$

$$\frac{d^2}{d\mu_0^2} H[f_M] = \frac{d\lambda_0}{d\mu_0} = -\frac{\begin{vmatrix} \mu_2 & \cdots & \mu_{M+1} \\ \vdots & \cdots & \vdots \\ \mu_{M+1} & \cdots & \mu_{2M} \end{vmatrix}}{\mid \Delta_{2M} \mid} = -\frac{1}{\mu_0 - \mu_0^{-(M)}} < 0.$$

Thus $H[f_M]$ is a concave differentiable function of μ_0. When $\mu_0 \to \mu_0^{-(M)}$ then $H[f_M] \to -\infty$, whilst at μ_0 it holds $H[f_M] > H[f]$, being $f_M(x)$ the maximum entropy density once assigned $(\mu_0, ..., \mu_M)$. Besides, when $M \to \infty$ then $\mu_0^{-(M)} \to \mu_0$. So the theorem is proved.

HAUSDORFF MOMENT PROBLEM VIA FRACTIONAL MOMENTS

Pierluigi Novi Inverardi[(1)], Alberto Petri[(2)], Giorgio Pontuale[(2)], Aldo Tagliani[(1)(*)]

[(1)] Faculty of Economics, Trento University, 38100 Trento, Italy.
[(2)] CNR, Istituto di Acustica "O.M. Corbino", 00133 Roma, Italy.
[(*)] Corresponding author:
Phone: +39-0461-882116, Fax:+39-0461-882124, E-mail: ataglian@cs.unitn.it

Abstract
We outline an efficient method for the reconstruction of a probability density function from the knowledge of its infinite sequence of ordinary moments. The approximate density is obtained resorting to maximum entropy technique, under the constraint of some fractional moments. The latter ones are obtained explicitly in terms of the infinite sequence of given ordinary moments. It is proved that the approximate density converges in entropy to the underlying density, so that it demonstrates to be useful for calculating expected values.

Key Words: Entropy, Fractional moments, Hankel matrix, Maximum Entropy, Moments.